DOCUMENT RESUME

ED 314 396 SP 031 867

AUTHOR Gilman, Muriel
TITLE Should Norm-referenced Standards or Criterion-referenced Standards Be Used in Evaluating Youth Fitness Tests?
PUB DATE 29 Sep 89
NOTE 8p.; Paper presented at the American College of Sports Medicine Northland Chapter Meeting (September 29, 1989).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Criterion Referenced Tests; Elementary Secondary Education; *Norm Referenced Tests; *Physical Fitness; *Physical Health; Test Selection; *Test Use

ABSTRACT

In this discussion of whether youth fitness tests should be evaluated with norm-referenced standards or criterion-referenced standards, it is pointed out the the suitability of the test depends upon the purpose for which the test is administered. A high level of physical fitness is linked with health, at least in adults, which suggests that physical fitness testing is done for the purpose of health screening and motivating students. For this purpose, criterion-referenced standards are more suitable. Norm-referenced standards are useful for comparing groups and selecting talented athletes, but do not necessarily link health and fitness. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. B. Gilman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED314396

"Should Norm-referenced standards or criterion-referenced standards be used in evaluating youth fitness tests?"

Presented at the American College of Sports Medicine Northland Chapter Meeting, September 29, 1989

Muriel Gilman, Ph.D.
Bemidji State University

The purpose of this paper is to answer the question of whether youth fitness tests should be evaluated with norm-referenced standards or criterion-referenced standards. As one asks this question, one must ask why is fitness testing being done? Is the purpose to perdict performance, to screen for health problems, to compare the schools performance to National Norms, to evaluate a program, to evaluate a student's progress, to educate students about the components of fitness, or to motivate students to become more physically fit? The purpose of testing will help in determining the best evaluation procedure.

Criterion-referenced standards involve setting some predetermined standard which carries normal risk of the development of a disease or increased risk of the disease if the criterion score is not met. Cholesterol screening is one of the most popular and visible examples of criterion-referenced standards. For example, a serum cholesterol less than 220 mg/dl carries normal risk while a level above 220 md/dl increases the risk of heart disease.

Traditionally, norm-referenced standards (NRS) have been used in the evaluation of fitness tests. This involves testing a large number of children and plotting the distribution of their

SP031867

scores on a particular fitness test item and then assigning the scores a percentile ranking. The percentile ranking is used to compare a child's score with the National Norm, or what may also be referred to as the reference population. Norm-referenced standards are helpful when comparisons between individuals are desired, as in selecting a team or identifying talented athletes. Well-skilled students are motivated by NRS evaluations, as they feel good about being told, for example, that they are in the 90th percentile. But, the student who does not score well is seldom motivated by being told that she/he scored in the 20th percentile. In fact, it probably just confirms what she/he already felt about his/her skill.

Another problem that may occur with NRS standards and motivation is the maturation rate of a child compared to the maturation of the reference population. With most NRS the same percentile ranking from year to year requires a "better" raw score, as most children can run longer or faster as they mature. If a child does not mature as rapidly as the average child she may run just as fast this year as she did last year but have a lower percentile ranking. Interpretation of such scores must be done with care so as not to allow discouragment.

Therefore, the NRS may be motivating for the fit child but imagine if you were unfit. You know that you have to take "those" fitness tests again, and when you get the results you are told that you scored at the 30th percentile for the 1 mile run. You tried hard, you even did some running over the summer,and

maybe you even improved a little but most of the children in the country still run faster than you. It is unlikely that such a child will want to continue exercising, after all she has been told that she is not very good at it.

Another problem with the NRS is the selection of the reference population. If the reference population was relatively fit the percentile ranking for a given raw score on would be lower, in other words the normal curve shift to the right. Whereas, if the reference population was less fit, the curve would shift to the left and the same raw score would yield a higher percentile ranking. In the past, the two most popular fitness tests, the Presidential Physical Fitness Award Program and the AAHPERD Health Related Physical Fitness Test have used different reference populations. This has resulted in different norms for common items. This could result in confusion for students if the school changed tests from year to year.

As the awareness of the relationship between physical fitness and health has increased the composition of physical fitness tests has changed. In the past, fitness tests were composed of motor fitness items, those skills for which we have an inherent ability and little improvement can be expected; and health related items, those which can be improved with training such as flexibility and cardiovascular endurance. The motor fitness items lend themselves to NRS because faster or further implies better performance of a skill. But, the newer fitness test batteries are mainly composed of health related fitness

items for which a particular level of achievement is related to improved health. This places increased importance on a criterion score which can predict health.

Criterion-referenced standards (CRS) are based on a score determined by experts which is thought to be sufficient physical fitness to prevent common degenerative diseases. This implies a relationship between health and fitness. Each test item and its respective criterion-referenced standard is based on its link with health, whether it is the risk of obesity, heart disease, low back pain or injury prevention. An unlimited number of students can achieve the criterion score whereas by the very nature of NRS only a limited number of students can achieve a high percentile ranking. In the AAHPERD Physical Best Program criterion referenced standards there are only a limited number of scores, allowing for some maturational effect. However, once maturation has occurred the criterion score changes very little, for example, 25 sit-ups is the CRS for girls from 9 years to 17 years. This allow a child to focus on a particular level of fitness and personal improvement becomes more important than comparisons between students. (Physical Best also allows a child to set his or her own goals and receive awards for achievement of those goals.)

It appears that the CRS are the best way of evaluating physical fitness tests at this time, but there are some problems with CRS. One is the lack of information about what level of fitness in childhood results in health in adulthood or old age.

Another is the lack of information about the relationship between childhood fitness and adult health. There is a need for prospective, longitudinal studies which look at childhood fitness levels and the incidence of various health problems in adulthood.

Dennison, et al. conducted a retrospective study on 453 young men who were 23-25 years at the time of the study. A comparison of their youth fitness test scores when they were 12 and 13 years old revealed that those who scored poorly on fitness tests as teenagers were less active as adults. So, even though we do not have data about the effect of childhood fitness on adult health per se, this study suggests what has seemed logical, that the less fit youth will be less active as an adult and research has revealed the importance of physical fitness for health in adults.

The answer to the question as to which type of standard should be used becomes more clear as one answers the question of, "Why is fitness testing done?". The change in fitness test batteries has sent the message that a high level of physical fitness is linked with health, at least in adults, which suggests that physical fitness testing is done for the purpose of health screening and motivating students to develop a higher level of fitness. Criterion-referenced standards can be motivating, educational, and useful for comparisons and selections, as well as health screening. Norm-referenced standards are useful for comparing groups and selecting talented athletes but do not necessarily link health and fitness. Therefore, CRS are the best

choice at this time. This conclusion is supported by the fact that several organizations involved in fitness and health have supported the use of CRS including AAHPERD, the American College of Sports Medicine and the American Academy of Pediatrics.

In summary, the 1980's have brought changes in fitness testing which have emphasized the link between fitness and health. This carries a message about what is important in fitness: cardiovascular endurance, muscular strength/endurance, body composition, and flexibility. As we enter the 1990's the documentation of the criterion scores will evolve so that as fitness scores are evaluated more confidence can be placed in the link between fitness and health in children.

BIBLIOGRAPHY

Blair, S.N., Falls, H.B., and pate, R.R. (1983) A New physical Fitness Test. The Physician and Sportsmedicine, 11(4):87-95.

Dennison, B.A., Straus, J.H., Mellitis, E.E. and Charney E. (1988) Childhood Physical Fitness Tests: Predictor of Adult Physical Activity Levels? Pediatrics, 82(3):324-330.

Going, S. and Williams, D. (1989) Understanding Fitness Standards. JOPERD, 60(aUG):34-38.

Opinion Statement on Physical Fitness in Children and Youth. (1988) Med. Sci. Sport Exerc., 20(4):422-423.
Physical Best

Physical Best. AAHPERD, 1900 Association Drive, Reston, VA, 22091.

Physical Fitness and the Schools. (1987) Pediatrics, 80(3):449-450.

Presidential Physical Fitness Award Program, Instructor's Guide. Department of Health and Human Services, The President's Council on Physical Fitness and Sports, Washington, DC, 20001.

Ross, J.G. and Pate, R.R. (1987) A Summary of Findings. JOPERD, 58(Nov-Dec):51-56.

Ross, J.G., Pate, R.R., Delpy, L.A., Gold, R.S. and Svilar, M. (1987) New Health-Related Fitness Norms. JOPERD, 58(Nov-Dec):66-70.

Whitehead, J.R. (1989) Fitness Assessment Results-Some Concepts and Analogies. JOPERD, 60(Aug):39-43.